IRIS - Archivio Istituzionale dell'Università degli Studi di Sassari

On the distinctiveness of the electricity load profile

Questa è la versione Pre print del seguente articolo:

Original
On the distinctiveness of the electricity load profile / Bicego, M., Farinelli, A., Grosso, E., Paolini, D., Ramchurn, S.D.. - In: PATTERN RECOGNITION. - ISSN 0031-3203. - 74:(2018), pp. 317-325. [10.1016/j.patcog.2017.09.039]

Availability:
This version is available at: 11388/198398 since: 2019-07-03T16:43:45Z

Publisher:

Published
DOI:10.1016/j.patcog.2017.09.039

Terms of use:

Chiunque può accedere liberamente al full text dei lavori resi disponibili come “Open Access”.

Publisher copyright

note finali coverpage

(Article begins on next page)

14 September 2026

On the distinctiveness of the electricity load profile

M. Bicego[a,*], A. Farinelli[a], E. Grosso[b], D. Paolini[b], S. D. Ramchurn[c]

[a] *University of Verona (Italy)*
[b] *University of Sassari (Italy)*
[c] *University of Southampton (UK)*

Abstract

The recent increasing availability of fine-grained electrical consumption data allows the exploitation of Pattern Recognition techniques to characterize and analyse the behaviour of energy customers. The Pattern Recognition analysis is typically performed at group level, i.e. with the aim of discovering, via clustering techniques, *groups of users* with a coherent behaviour – this being useful, for example, for targeted pricing or collective energy purchasing. In this paper we took a step forward along this direction, investigating the possibility of discriminating the behaviours of *single users* – i.e., in a biometrics sense. This aspect has not been properly addressed and would pave the way to crucial operations, such as the derivation of alternative advertising schemes based on behavioural targeting. To investigate the uniqueness of the load profiles (i.e. the daily consumption of electrical energy), in our study we used the raw data (the original energy consumption time series) as well as different types of features such as frequency coefficients and normalized load shape indexes, together with various classification schemes. Results obtained on two real world datasets suggest that the load profile does contain significant distinctive information about the single user.

Keywords: Energy market, load profile, biometrics, classification, pre-processing

*Corresponding author

1. Introduction

Over the past two decades, power and energy systems have been experiencing a huge transformation, due to the increase of importance of renewable energy sources such as solar, hydroelectric and wind power. In this perspective, the balancing of power sources and consumer demand becomes a serious challenge that cannot totally rely on local production and energy storage systems, but rather requires non isolated grids and intelligent reversal of the load flows, following customer needs. This drastic change opens new and challenging problems for intelligent control systems which must face a number of new interesting issues: in this sense Pattern Recognition tools [1] may be of paramount importance, being able to provide solutions to problems such as forecasting of energy prices, optimal dispatching, consumer segmentation, and energy demand allocation [2, 3, 4, 5]. In particular, the availability of fine-grained electrical consumption data (due to the recent large scale deployment of intelligent metering infrastructures), coupled with an increasing and worldwide energy market liberalization, results in a growing interest in discovering and categorizing *groups of users* which share similar behaviours. This is usually done via clustering of the so called user *load profiles*, i.e., the users' consumption of electrical energy over a given period, as measured by the so called Advanced Metering Systems (AMS). Interestingly, this has to do with models taking into account the dynamics and the magnitude of the consumption and the ability to capture in the models exogenous factors (type of appliances, insulation) or context factors (occupancy, weather, seasons, holidays). With reference to figure 1, a typical processing scheme includes the following steps [6]:

- temporal aggregation;
- context filtering;
- metadata generation;
- data analysis.

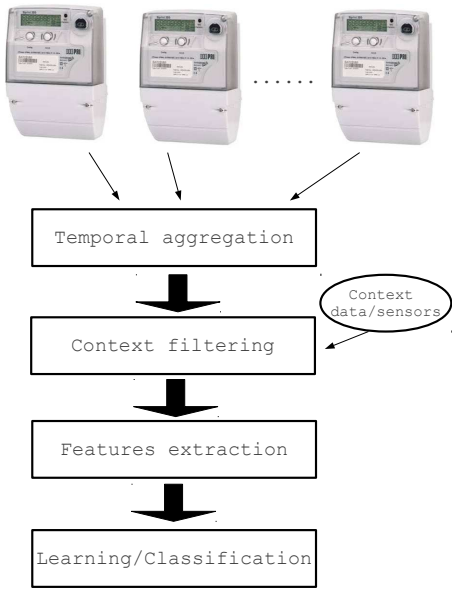

Figure 1: Typical processing scheme of data collected from AMS

In this scheme, temporal aggregation is used to define the temporal granularity of the data consumption collection (hourly, daily, etc.) while the context filtering stage takes into account specific factors such as holidays, seasons and temperatures. Metadata generation is probably the most critical step in the proposed processing scheme. In fact, starting from a coherent set of temporal measures (load profiles), a number of quantitative descriptors can be derived; in the literature, these descriptors are often denoted as feature functions [6] because they act on the load profiles transforming the original time representation into a more compact or more discriminative representation. Examples of quantitative descriptors are the load factor and the night/lunch impact [7, 8]; the Fast Fourier Transform is another example of data manipulation giving evidence to the content, in the frequency domain, of the original time representation.

The final analysis step is typically devoted to the clustering of the various load profiles, in order to detect coherent groups of users. This task can be very dificult due the number of groups which is generally unknown and the number

of users that can be very high in real applications.

Different approaches have been proposed to face this problem: for example, in [7] authors propose a framework to characterize groups of users based on simple load descriptors. They also prove the robustness of the proposed method with respect to missing data and outliers. Carpaneto and colleagues [9] propose a scheme based on the frequency domain. In contrast, other approaches [10, 11] consider each load profile as time sequence of load measurements and apply various unsupervised learning techniques for clustering (Self Organizing Maps, K-means and Hidden Markov Models among others). A good overview of applicable pattern recognition tools is given in [12]; this paper also includes a detailed description of most interesting clustering methods and proposes several consistency measures adequate to evaluate the performance of these methods. Another review is given in [13], discussing in particular how the number of categories can vary depending on locations and type of loads (public, industrial, residential). A deep investigation of clustering methods applied to the domestic sector in Ireland has been recently presented in [14]; authors consider few profile categories and a customer is essentially defined by a vector of likelihood coefficients, showing the statistical association of the customer to each of the profile groups defined.

Even though the above work represent an impressive progress in this area, there is an urgent need for advanced analysis of energy usage data. For example energy companies are becoming more and more interested in targeted advertisement, personal tarification [15] or even in detecting frauds [16] and changes in the composition of the group of people living in a given house. Analysts have begun to use these data for different goals, such as for example the optimal allocation of the energy flows and the reduction of purchase prices, or to help retailers designing new pricing models for implementing more accurate demand and supply profiles [10, 8, 11]. For all these applications, methods which work *at group level* are not enough, since the characterization and discrimination should be done *at the user level*: in other words, there is a need for automatic systems able to characterize and discriminate *every user* related to a single metering

system: this crucial aspect has never been investigated in the literature, and represents the main goal of this paper. In particular, starting from some preliminary and encouraging results [17], this paper investigates different types of metadata and classification schemes to understand if an answer exists to the following key question: *does every single user have a unique behaviour when consuming electrical energy?* or, in different terms, *can the electrical energy consumption related to a single AMS be considered as a distinctive behavioral trait?* As better explained in the following, an answer to this question may open the possibility of devising novel targeting strategies and at the same time it would spur an important discussion on important privacy issues.

To answer the above question, in this paper we develop a *classification system* to identify a specific user (or, more precisely, a specific AMS) among several users, on the base of the electrical consumptions over a given period of time (i.e., a load profile). We investigate different metadata characterizing load profiles, including raw measurements, frequency characterizations and typical load shape indexes. We also investigate two classification schemes: the former is based on the classical Nearest Neighbor rule (i.e., it assigns an unknown object to the class of its nearest neighbor); the second scheme follows the classical Bayesian classification [1], based on Hidden Markov Models (HMM – [18]). This probabilistic approach has been widely used to characterize sequential data, and has been recently applied to the problem of clustering load profiles (in particular to characterise relationships between consumers' preferences or behaviors and electricity consumption [11]). The empirical evaluation is based on two databases composed of real load profiles, collected in the UK and in Portugal from several hundreds of metering systems. The system is trained on a known set of profiles, and tested on an hold out set. Our classification results suggest that the energy load profile does indeed contain user-specific discriminative information.

The rest of the paper is organized as follows: Section 2 details the problem of personal tarification and behavioral targeting, unfolding the complexity of the problem and the potential benefits of a behavioral analysis, also from an economic point of view. Section 3 presents the proposed approach, also in

relation to related works, detailing both the choice of metadata and the proposed classification scheme. The empirical evaluation of the proposed approach is given in Section 4, while Section 5 concludes the paper.

2. Personal targeting for the energy market

In this section we provide some considerations on the impact that the distinctiveness of the user load profiles may have on the energy market. In particular, we are convinced that the behavioural peculiarities of the load profile may lead to the so-called behavioural targeting (BT) [19] in the energy market. BT is a kind of advertising that is based on the analysis of the peculiar and distinctive behaviour a user shows in a specific context. Starting from the seminal analysis of Grossman and Shapiro in 1984 [19], the interest of economics literature on this field has grown significantly. From a market perspective, it should be also noted that advertising using behavioural targeting is becoming an important industry: eMarketer estimated that online advertisers spent more than $1.3 billion in targeted advertising in 2011, and the figure was expected to rise to more than $2.6 billion in 2014 [20]. Usually, when speaking about BT, we refer to media frameworks. Platforms (such as Google), have access to technologies allowing to gather information on the behaviour of platform users, making it possible to customize consumers (see [21]). Advertisers may submit bids that depend, for instance, on the correspondence between the website's content and the advertisement, but also on data about the location of the consumer (obtained through the Internet Protocol IP address). This kind of advertisement makes sense also because the behaviour of people in Internet contains distinctive traits (for example, it has been shown that it is possible to distinguish between different users on the basis of the browsing histories [22] or even the way they marked as favourite Flickr pictures [23].)

The introduction of new technology, such as modern AMS, in the electricity market allows to collect data related to the possibly distinctive energy consumption style – this typically results in group-based targeting/pricing. We

are convinced that this would be pushed even more, by effectively applying behavioural targeting also in this non-internet-based system; clearly, as a crucial starting point, the energy load profile should contain distinctive behavioural information – and this represents the main scope of the present work.

From the economic point of view, as pointed out by [21], the disclosure of this behavioural information can improve the match between advertisers and consumers; however it remains crucial to reply to the following two questions: i) is it profitable for the advertisers? ii) is it good for the consumers? The first reply is quite simple if we observe the data. As reported above, the eMarketer estimated that online advertisers spent more than $1.3 billion in targeted advertising in 2011 (rising to more than $2.6 billion in 2014 [20]).

The situation is quite different for the consumers. On the positive side, the advertiser can condition their contract proposal on information about consumers, the better the information delivered to the advertiser, the more focused the proposal. In short, the matching between the supply (advertiser) and the demand side (consumers) will be efficient. On the negative side, since good matches correspond to higher marginal revenues for advertisers, it can be shown that disclosure of personal information leads to higher prices for consumers. The fact that disclosure can lead to higher prices, for [21], comes from the observation that the firms can condition their proposal on consumers' characteristics: more specifically, the disclosure of information leads to a situation in which firms expect their ads to reach only the consumers with a low price-elasticity of demand (the good matches). Firms then rationally set higher prices ex ante. Moreover, as outlined by [24], there is the possibility that higher prices can derived from the higher costs of the targeting.

Whether positive effects compensate negative effects remains an open issue, to be investigated in a more economic oriented future work. However, trying to understand whether electricity load profiles can be used to identify user-specific behaviors is a key question and is the main focus of this work.

3. The classification scheme

In this paper we aim at discriminating single users based on their load profile and to this end we focus on a classification problem. As mentioned in the introduction, several works in the literature consider the analysis of users in the energy domain, proposing a wide spectrum of approaches, various types of metadata and several techniques [14, 6]. However, all these approaches focus on the clustering problem, i.e. on the aggregation of users in few groups. Moreover, several studies reveal that the consistency of user aggregation can vary significantly depending on various characteristics (such as season, features of the house and behaviors of users [6]). This interesting analysis suggests that the energy load profile might be a powerful trait to discriminate users. Specifically, in this section we detail our methodology for user identification. The starting point is the "electricity load profile", that is the recorded energy consumption of each user throughout the 24 hours of the day (typically every 15-30 minutes). According to this definition, a single profile is a vector of measurements and a single user is characterized by a set of profiles, one for each day. Figure 2 gives a visual representation of such load profiles for different users and different days (data from the EnergyUK dataset – see the experimental part).

As previously stated, the load profile can be characterized by considering various metadata. In particular we employ here three different representations based on time, frequency and load shape indexes [7, 25]. Given these metadata, two different classification schemes are proposed: the former based on the nearest neighbor approach and the second based on probabilistic HMM. Since the application of HMM is specific to sequential data, only time representation is used in the second classification scheme.

3.1. Metadata definition

Time. The full load profile is used, namely the T-dimensional vector returned by the acquisition device. Note that this choice is quite common in clustering

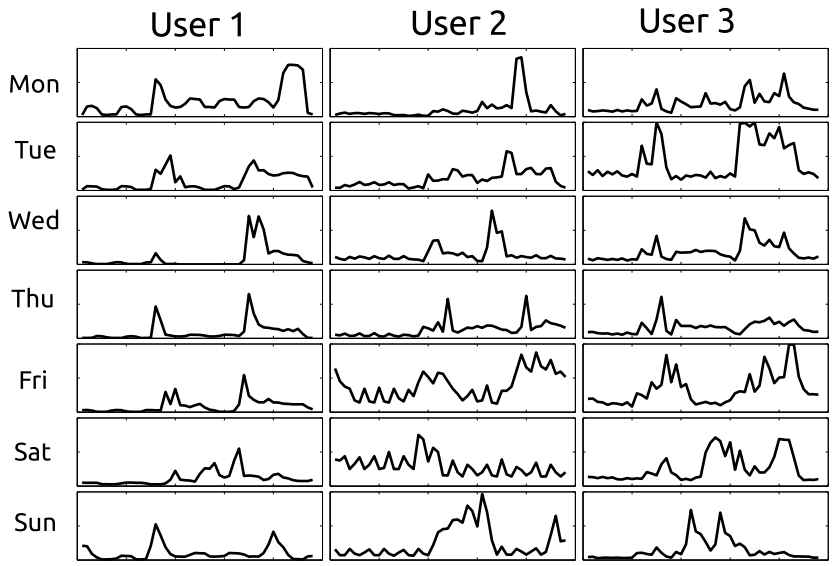

Figure 2: Examples of load profiles.

energy profiles – e.g. [10, 11].[1] The signals are not normalized (for example by applying z-score normalization), so to maintain all possible information on the energy of the signal (the absolute scale of the consumption profile) which might be useful to differentiate among users[2].

Frequency. In order to investigate how the frequential content may be helpful to discriminate among users, Fast Fourier Transform is applied to the load profile. After a careful evaluation of the frequency content of the profiles, only the first 10 coefficients have been retained.

Normalized load shape indexes. These feature functions derive from the load profile a compact set of values (indexes). These indexes are effective in the characterization of groups of users, as recently shown by Figueireido and

[1] Note that a small level of Gaussian smoothing is applied in order to remove some fluctuations due to the sampling intervals.

[2] Actually, different types of normalization, including the z-score normalization have been evaluated; results are not shown here but the normalization schemes decreased the performance of the classification approach in all the experiments.

colleagues [7, 25]. Since they are widely used in various studies, shape indexes are included in the present work: in particular, a selection of three indexes suggested in [7] is applied:

1. **Load Factor**:
$$LF = \frac{P_{av,day}}{P_{max,day}} \tag{1}$$
where $P_{av,day}$ is the average consumption computed throughout the whole day, and $P_{max,day}$ represents the highest recorded peak of consumption.

2. **Night Impact**:
$$NI = \frac{1}{3}\frac{P_{av,night}}{P_{av,day}} \tag{2}$$
where the average consumption during the night, $P_{av,night}$ is related to the data recorded from 11pm to 7am (8 hours in total).

3. **Lunch Impact**:
$$LI = \frac{1}{8}\frac{P_{av,lunch}}{P_{av,day}} \tag{3}$$
where the average of the profile during the lunch time $P_{av,lunch}$ is assumed to last from 11.30am to 2.30pm (3 hours in total).

The three shape indexes above are computed for every load profile; a single day of a given user is thus characterized as a 3-dimensional feature vector.

3.2. Classification

The goal of the proposed study is to investigate the uniqueness of the electric load profile of a given user. To do that, a key issue is the design of a classification scheme, that is a method to assign a given unknown profile into one over a set of predetermined users. In what follows we detail the two classification schemes adopted (both are basic, well known approaches) and their application for load profile classification.

Nearest Neighbor Scheme. In the Nearest Neighbor scheme, the classifier assigns an object to the class of its nearest neighbor; the definition of a suitable distance measure (can be either based on "similarity" or "dissimilarity" features) is therefore crucial. The distance metrics adopted in the present work

are introduced in the following, together with the motivations behind these choices. Note that the L1 and the L2 norms (the Manhattan and the Euclidean distances, respectively) are the first distance metrics introduced. In fact these metrics are very common, widely applied for Pattern Recognition purposes, and also specifically applied to clustering problems in the energy domain (see for instance [10]). Given two profile representations $\mathbf{p} = p_1, ..., p_T$ and $\mathbf{q} = q_1, ..., q_T$, L1 and the L2 norms are defined as follows:

L1 norm (Manhattan distance):

$$L1(\mathbf{p}, \mathbf{q}) = \sum_i |p_i - q_i| \tag{4}$$

L2 norm (Euclidean distance):

$$L2(\mathbf{p}, \mathbf{q}) = \sqrt{\sum_i (p_i - q_i)^2} \tag{5}$$

It is worth noting that L1 and L2 norms make sense for all the considered metadata; on the other hand we also investigate several additional metrics specifically designed for the time signal representation and based on the concept of signal correlation, which is a standard and widely used method to compare time series [26].

Given two profiles $\mathbf{p} = p_1, ..., p_T$ and $\mathbf{q} = q_1, ..., q_T$, standard Zero Lag Cross Correlation (ZLCC) is defined as:

Zero Lag Cross Correlation:

$$CC_0(\mathbf{p}, \mathbf{q}) = \frac{\sum_i (p_i \cdot q_i)}{\|\mathbf{p}\| \|\mathbf{q}\|} \tag{6}$$

This is the first additional metric adopted. The second metric investigated, again based on signal correlation, takes into account the fact that activities characterizing a specific user in different days may not be completely overlapped but there can be a small time lag; for example this could happen if an early departure for work leads to anticipate a number of activities with respect to a regular day. In order to capture this behaviour, cross correlation can be computed by allowing some time steps of lag, and retaining at the end the

maximum of the correlations. Following this rationale, two lag-based measures have been adopted:

Max (1-lag) Cross Correlation:

$$CC_1(\mathbf{p},\mathbf{q}) = \max_{m\in\{-1,0,1\}} \frac{\sum_i (p_i \cdot (q_i + m))}{\|\mathbf{p}\|\|\mathbf{q}\|} \tag{7}$$

Max (2-lag) Cross Correlation:

$$CC_2(\mathbf{p},\mathbf{q}) = \max_{m\in\{-2,-1,0,1,2\}} \frac{\sum_i (p_i \cdot (q_i + m))}{\|\mathbf{p}\|\|\mathbf{q}\|} \tag{8}$$

Finally, we also investigated the fact the lags displacement in the daily activities can vary depending on the kind of activity carried on; for example the way people watch the TV or have a dinner can be similar but watching the TV and having a dinner can take place at different hours in different days. In order to capture this behaviour we repeated the computation of the Max 2-lag cross correlation (as defined above) for small overlapping windows lasting 4 hours and overlapping for two hours. This strategy allows to best align consumptions related to different parts of the day, taking at the end the mean or the max computed correlation value. More precisely, defining as z the total number of overlapped sub-windows $\mathbf{w}_i(\cdot)$ extracted from a given profile and lasting exactly 4 hours, two additional measures have been defined as:

Mean Windows Max (2-lag) Cross Correlation:

$$CC_{MeanW2}(\mathbf{p},\mathbf{q}) = \frac{1}{z}\sum_{i=1}^{z} CC_2(\mathbf{w}_i(\mathbf{p}),\mathbf{w}_i(\mathbf{q})) \tag{9}$$

Max Windows Max (2-lag) Cross Correlation:

$$CC_{MaxW2}(\mathbf{p},\mathbf{q}) = \max_{i\in\{1,\ldots,z\}} CC_2(\mathbf{w}_i(\mathbf{p}),\mathbf{w}_i(\mathbf{q})) \tag{10}$$

Hidden Markov Model-based Bayesian classification. This scheme revives the classical Bayesian decision scheme [1], which assigns an unknown pattern to the class which shows the highest posterior probability, or, assuming equiprobable classes, the highest class conditional probability. In our case, class conditional probabilities are modelled using Hidden Markov Models [18], a probabilistic technique whose effectiveness has been shown in various recognition

scenarios. In few words, a discrete-time first order HMM [18] is a probabilistic model that describes a stochastic sequence $\boldsymbol{O} = (O_1, O_2, \ldots, O_T)$ as being an indirect observation of a hidden Markovian random sequence of states $\boldsymbol{Q} = (Q_1, Q_2, \ldots, Q_T)$, where $Q_t \in \{1, 2, \ldots, N\}$ (the set of states), for $t = 1, ..., T$. Associate to each state there is a probability function, called emission probability function, which describes the probability of emitting a given symbol from such state. A HMM is then completely specified by a set of parameters $\boldsymbol{\lambda} = \{\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{\pi}\}$ where $\boldsymbol{A} = (a_{ij})$ is the transition matrix (a_{ij} is the probability of passing from state i to state j), $\boldsymbol{\pi} = (\pi_i)$ is the initial state probability distribution (π_i is the probability that system starts from state i) and $\boldsymbol{B} = (\boldsymbol{b}_i)$ is the set of emission probability functions, i.e. $\boldsymbol{b}_i(o)$ is the probability of emitting the symbol o when the system in state i. In the considered case, the observations are continuous, therefore we assume that each $\boldsymbol{b}_i$ is a Gaussian probability density function.

Given a set of sequences $\{\boldsymbol{o}^{(i)}\}$, the training of the model permits to estimate the best parameters $\boldsymbol{\lambda} = \{\boldsymbol{A}, \boldsymbol{B}, \boldsymbol{\pi}\}$, i.e. the parameters that maximize the probability $P(\{\boldsymbol{o}^{(i)}\}|\boldsymbol{\lambda})$. This step if typically performed by using the Baum-Welch re-estimation technique [18]. Given a sequence $\boldsymbol{o}$, it is possible to evaluate how well this sequence is modelled by the HMM, using a procedure called *evaluation*. In particular, using a procedure called *forward-backward procedure* [18] it is possible to estimate the log probability $\log P(\boldsymbol{o}|\boldsymbol{\lambda})$, given a model $\boldsymbol{\lambda}$ and the sequence $\boldsymbol{o}$ to be evaluated.

To summarize, given a C-class problem, our Bayesian classification scheme is realized in the following way: for every class c, corresponding to a given user, a HMM $\boldsymbol{\lambda}_c$ is trained; to this purpose only the training sequences (profiles) belonging to such class are used, obtaining as a result the set of C models $\boldsymbol{\lambda}_1, ..., \boldsymbol{\lambda}_C$. In the subsequent testing phase, an unknown load profile $\boldsymbol{o} = (o_1, ..., o_T)$ is assigned to the class whose model shows the highest likelihood (note that to each class is assigned the same prior probability).

4. Experimental evaluation

This section contains the empirical evaluation for our approach. We first describe our empirical methodology and then present a wide set of results, investigating different representation schemes and different versions of the recognition approach. Subsequently, we introduce an analysis on how well the introduced schemes scale with the size of the dataset. Further, we propose some considerations and experiments on how it is possible to aggregate more days to characterize a given user. Finally, we present some considerations on how sensible these results are with respect to the different seasons.

4.1. Empirical Methodology

The data employed in our experiments derive from two sources, both related to settings typically employed in the domain of collective energy purchasing, where the goal is to find group of energy consumers in order to purchase electricity at more convenient fares [27, 28].

More precisely, in the first dataset (called *EnergyUK*), a load profile contains the recording of the consumption of energy in a given household at fixed intervals (half hour); data have been recorded over a period of one month, in 2009, for several houses in UK. In Fig. 2 we provide some examples of load profiles, covering different customers in different days. The second dataset, which we call *EnergyPT*, contains again daily consumption, but now recorded every 15 minutes (for a total of 96 values per day), related to several different Portuguese clients recorded during the period 2011-2014 [29].

Most of the tests have been performed by using the first dataset (*EnergyUK*): actually, the fact that all the load profiles have been gathered during the same month allows to investigate only person-specific characteristic behaviours, leaving aside the possible presence of discriminant traits due to seasonality. However, to investigate a larger temporal domain, we also performed some experiments on the *EnergyPT* dataset, in order to understand how difficult is to classify profiles while changing seasons.

As in any classification test, it is important to keep separated the data used for training the system from the data used to test it. Here, for the (*EnergyUK*) dataset, we employed as training set the first 2 weeks of the month, using the remaining days as testing set. For what concerns the *EnergyPT* dataset, – as it will be clearer later – we employed signals derived from one season to train the system, whereas testing is performed with signals deriving from the other seasons. Our analysis employs the so-called Cumulative Match Curve (CMC); this represents a widely used performance indicator, typically employed when testing behavioural biometrics [30] and is more informative than the error of the classification scheme. To compute such measure, for every testing load profile a ranking of the identities matched by the system is determined: clearly, the top matched identity (i.e. the identity for which the matching score is maximum) represents the class assigned to such testing profile. Given the ranking, the CMC at abscissa k indicates the percentage of times at which the correct identity is found within the top k matches: k is spanned along all possible values (from 1 to the number of classes). In order to summarize the CMC, typically the normalized Area Under the Curve (nAUC) is measured. This parameter, similarly to what is done for the ROC curve, represents an useful summarizing measure: the higher this value, the better the recognition capabilities of the investigated system.

For what concerns Hidden Markov Models, in all the experiments the training has been carried out using the classical Baum-Welch procedure [18]; this iterative procedure stops when the log likelihood reaches a stable value. The Baum-Welch procedure performs a local optimization, therefore it is highly sensible to the initialization of the parameters. Following a classical approach employed in different applications dealing with Continuous Gaussian HMM, the parameters have been initialized via clustering, employing a Gaussian Mixture model. A free parameter of Hidden Markov Models is represented by the number of states, which drives the complexity of the model. Here we adopted an automatic scheme described in [31], which determines the best number using only the training set.

4.2. First Experiment: comparing different features and classification schemes

In the first group of experiments we focused on a set composed by 100 users extracted from the *EnergyUK* dataset, which we analysed by using various metadata and various versions of the 2 classification approaches. In this case, the main objective is to determine the best configurations, usable in the subsequent tests. To smooth the raw profile (when applied), we employed a Gaussian filtering – sigma varies between [0.6 - 2.2]. The nAUC values, for all cases, are reported in table 1. In the best case, the nAUC is 0.927: considering that the signal is a behavioural trait, this represents a reasonably high value – please compare with [32], which analyses more established traits (like voice or gait). Among the two classification schemes, the Bayesian approach seems to perform best, being able to capture, via the learning phase, the unique characteristics of every subject. Nonetheless, also the NN scheme is reasonably accurate, except when used with the correlation measures, which probably are too flexible. For what concerns metadata, it seems evident that the Load Shape Indexes, a good choice for general data-mining [7, 25], is not sufficient to capture the differences between subjects.

4.3. Second experiment: scalability

Here we study the capabilities of the presented strategies to scale in a reasonable way while increasing the number of the considered subjects. To conduct the experiments we selected the two best schemes, as determined in the previous section ("Smooth TR + L1 + NN" and "TR + HMM"), evaluating them with 200, 300 and 400 users of the *EnergyUK* dataset. Obtained results are presented in Table 2 – where, for sake of clarity, we presented also the values obtained with 100 users: interestingly, the performances do not drop when increasing the number of subjects considered.

4.4. Third experiment: aggregating several days in the load profile

In this section we investigate the possible effects of characterizing a subject by aggregating more days: actually, it is possible that the consumption of energy

Table 1: (*EnergyUK* dataset): Results with 100 users: "TR" stands for Time Representation, "FR" for Frequency Representation, 'LSI' for Load Shape Indexes

Nearest Neighbor Schemes		
Metadata	Classifier	nAUC
TR	L1 + NN	0.855
TR	L2 + NN	0.798
TR	CC_0 + NN	0.725
TR	CC_1 + NN	0.752
TR	CC_2 + NN	0.755
TR	CC_{MeanW2} + NN	0.780
TR	CC_{MaxW2} + NN	0.699
Smooth TR	L1 + NN	0.868
Smooth TR	L2 + NN	0.839
Smooth TR	CC_0+ NN	0.759
Smooth TR	CC_1+ NN	0.766
Smooth TR	CC_2+ NN	0.763
Smooth TR	CC_{MeanW2}+ NN	0.781
Smooth TR	CC_{MaxW2}+ NN	0.749
FR	L1 + NN	0.848
FR	L2 + NN	0.845
LSI	L1 + NN	0.722
LSI	L2 + NN	0.722
HMM schemes		
Metadata	Classifier	nAUC
TR	HMM + Bayes Rule	0.927
Smooth TR	HMM + Bayes Rule	0.905

varies according to the different day of the week (e.g. working days vs weekend). To do that, we consider as user signature 2, 3,..., 7 consecutive days. These profiles are aggregated by averaging (resulting again in a signature of length

Table 2: (*EnergyUK* dataset): Scalability results

Classes	100	200	300	400
NN approach	0.868	0.866	0.860	0.860
HMM approach	0.927	0.931	0.932	0.929

T) or by concatenation (obtaining a longer signature). In the former case we possibly remove noise (reducing variation inside the class), while in the latter we can have a larger set of (maybe noisy) data. Experiments were conducted using 100 and 200 subjects of the *EnergyUK* dataset, again with the "Smooth TR + L1 + NN" and "TR + HMM" selected in the previous sections. Table 3 presents the obtained nAUC.

Table 3: (*EnergyUK* dataset): Enriching the scope. "NN' and 'HMM' represent the two approaches experimented – see the text – 'Av' stands for Average, 'Co' stands for Concatenation.

Days	1	2	3	4	5	6	7
NN Av.	0.868	0.893	0.904	0.921	0.927	0.939	0.941
NN Co.	0.868	0.876	0.872	0.886	0.875	0.872	0.892
HMM Av.	0.927	0.906	0.901	0.893	0.905	0.900	0.910
HMM Co.	0.927	0.945	0.953	0.957	0.960	0.971	0.963

100 subjects

Days	1	2	3	4	5	6	7
NN Av.	0.866	0.888	0.901	0.915	0.920	0.937	0.933
NN Co.	0.866	0.872	0.869	0.878	0.866	0.875	0.877
HMM Av.	0.931	0.905	0.905	0.897	0.902	0.904	0.901
HMM Co.	0.931	0.948	0.960	0.965	0.967	0.972	0.970

200 subjects

In general, aggregating more days is beneficial for both approaches, with the performances raising from 0.868 to 0.941 for the Nearest Neighbor approach

and from 0.927 to 0.971 for the HMM approach. The former scheme gains more with signal averaging, whereas the latter prefers signal concatenation. This is somehow expected: the averaging operation produces more robust instances, which are essential for the NN scheme, based on pairwise comparisons; at the same time, however, the number of training instances is decreased, this makes the HMMs training less robust, being affected by the lower cardinality of the training set.

As a final picture of the performances for this dataset (EnergyUK), we present in Fig. 3 the CMC curves for the two best situations: NN (with averaging) and HMM (with concatenation). The figure clearly confirms the potentialities of our suggestion (i.e. that the user's load profile contains discriminative information): in the HMM case, by employing a single day, the correct identity of a given user can be identified within the top 20 answers in the 89.4% of the cases. When using more days, such percentage increases even more (94.5% with 7 days). Please note that a random classifier would present a 20% recognition rate. Similar conclusions can be drawn also in the case of 200 subjects (where the random classifier would obtain 10%).

4.5. Fourth experiment: considering seasons

In this section we investigate the robustness of the proposed approach to changes in seasons: in particular, we investigate the capability of the system to recognize a load profile collected in a specific season using a system trained with load profiles gathered in a different season. To do so we employed the signals of 100 users of the *EnergyPT* dataset, recorded during the year 2012. We split the signals in four sets, each one (approximately) related to a season (January, February and March for Winter, April, May and June for Summer, ad so on). We then train the system using a given season (e.g. Winter), testing it using another season taken from the the remaining ones (e.g. Spring, Summer and Autumn). We repeated the experiments testing all seasons, again with the "Smooth TR + L1 + NN" and "TR + HMM" selected in the previous sections – here we used a single day for the load profile. Table 4 presents the obtained

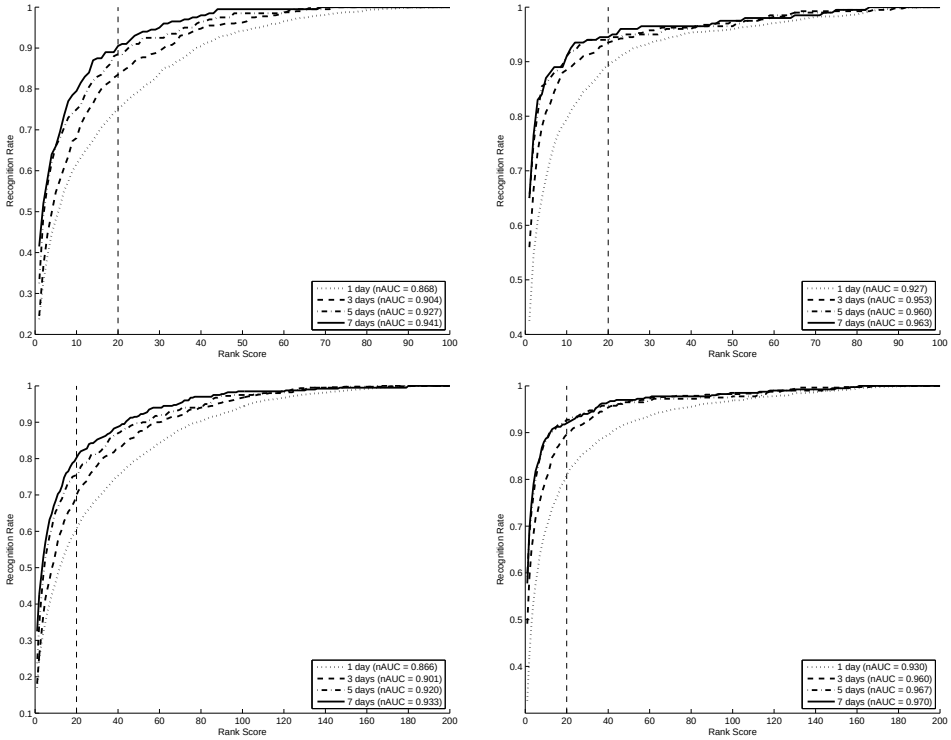

Figure 3: (*EnergyUK* dataset): CMC curves, for 100 (top) and 200 (bottom) subjects. The 20-rank rate is represented by a vertical dotted line. Left column is for the Nearest Neighbor, right column for HMM.

nAUC.

Interestingly, the recognition rates are very high: the system does not suffer too much for the change in the season, but it is able to capture the true traits of a given load profiles. Reasonably, the worst performances are obtained for winter/summer pairs. It is also interesting to note that the Nearest Neighbor technique outperforms the HMM-based scheme (whereas for the EnergyUK dataset we got the opposite behaviour). Probably, since here the temporal span of the training set is larger (3 month with respect to 15 days), the local NN, which takes the decision based on few profiles, is more suitable than the global HMM, which extracts a single models from all the profiles – which can be some-

Table 4: (*EnergyPT* dataset): Test on the different seasons, for (top) the HMM scheme and (bottom) the NN scheme.

	Testing season			
Training Season	Winter	Spring	Summer	Autumn
Winter	-	0.950	0.889	0.950
Spring	0.959	-	0.956	0.951
Summer	0.908	0.948	-	0.922
Autumn	0.965	0.959	0.919	-

Hidden Markov Model scheme

	Testing season			
Training Season	Winter	Spring	Summer	Autumn
Winter	-	0.985	0.955	0.974
Spring	0.989	-	0.992	0.974
Summer	0.962	0.983	-	0.959
Autumn	0.986	0.985	0.975	-

Nearest Neighbor scheme

how different, spanning three months.

In any case, also this experiment clearly confirms the potentials of using load profiles to discriminate users: in the NN case, for the pair Spring/Summer, the correct identity of a given user can be identified within the top 20 answers in the 99.65% of the cases, and within the top 5 in the 94.76% of the cases.

5. Conclusions

The deployment of AMS results in a large amount of fine-grained data that can provide crucial information for the energy distribution process. In this paper we focus on the discriminative power of energy load profile and specifically, we investigate whether the electricity load profile can provide distinctive information, to be exploited by intelligent control systems for advanced tasks such as

forecasting, balancing of energy loads, monitoring and optimization of consumptions as well as targeted initiatives. Our work provides empirical evidence that confirms the uniqueness of the behaviours of the users and proposes a practical classification scheme that allows to detect such unique patterns. This represents a preliminary viable tool that can be used by energy market societies to perform advanced analysis of the energy domain.

References

[1] R. Duda, P. Hart, D. Stork, Pattern Classification, 2nd Edition, John Wiley and Sons, 2001.

[2] J.-Y. Boivin, Demand side management – the role of the power utility, Pattern Recognition 28 (10) (1995) 1493–1497.

[3] R. Pino, J. Parreno, A. Gomez, P. Priore, Forecasting next-day price of electricity in the spanish energy market using artificial neural networks, Engineering Applications of Artificial Intelligence 21 (1) (2008) 53 – 62.

[4] A. Chrysopoulos, C. Diou, A. Symeonidis, P. Mitkas, Bottom-up modeling of small-scale energy consumers for effective demand response applications, Engineering Applications of Artificial Intelligence 35 (2014) 299 – 315.

[5] L. Marzio, A new utility-user interface for a qualified energy consumption, Pattern Recognition 28 (10) (1995) 1507–1515.

[6] T. Wijaya, T. Ganu, D. Chakraborty, K. Aberer, D. Seetharam, Consumer segmentation and knowledge extraction from smart meter and survey data, in: SIAM International Conference on Data Mining (SDM14), 2014, pp. 226–234.

[7] V. Figueiredo, F. Rodrigues, Z. Vale, J. Gouveia, An electric energy consumer characterization framework based on data mining techniques, IEEE Transactions on Power Systems 20 (2) (2005) 596–602.

[8] G. Chicco, R. Napoli, P. Postolache, M. Scutariu, C. Toader, Customer characterization options for improving the tariff offer, Power Systems, IEEE Transactions on 18 (1) (2003) 381–387.

[9] E. Carpaneto, G. Chicco, R. Napoli, M. Scutariu, Electricity customer classification using frequency–domain load pattern data, International Journal of Electrical Power & Energy Systems 28 (1) (2006) 13–20.

[10] T. Zhang, G. Zhang, J. Lu, X. Feng, W. Yang, A new index and classification approach for load pattern analysis of large electricity customers, IEEE Transactions on Power Systems 27 (1) (2012) 153–160.

[11] A. Albert, R. Rajagopal, Smart meter driven segmentation: What your consumption says about you, IEEE Transactions on Power Systems 28 (4) (2013) 4019–4030.

[12] G. Tsekouras, P. Kotoulas, C. Tsirekis, E. Dialynas, N. Hatziargyriou, A pattern recognition methodology for evaluation of load profiles and typical days of large electricity customers, Electric Power Systems Research 78 (9) (2008) 1494–1510.

[13] I. Prahastono, D. King, C. S. Özveren, A review of electricity load profile classification methods, in: Universities Power Engineering Conference, 2007. UPEC 2007. 42nd International, IEEE, 2007, pp. 1187–1191.

[14] F. McLoughlin, A. Duffy, M. Conlon, A clustering approach to domestic electricity load profile characterisation using smart metering data, Applied energy 141 (2015) 190–199.

[15] T. Krishnamurti, D. Schwartz, A. Davis, B. Fischhoff, W. B. de Bruin, L. Lave, J. Wang, Preparing for smart grid technologies: A behavioral decision research approach to understanding consumer expectations about smart meters, Energy Policy 41 (2012) 790–797.

[16] H. Khurana, M. Hadley, N. Lu, D. Frincke, Smart-grid security issues, IEEE Security & Privacy 1 (2010) 81–85.

[17] M. Bicego, F. Recchia, A. Farinelli, S. Ramchurn, E. Grosso, Behavioural biometrics using electricity load profiles, in: Proc. Int. Conf. on Pattern Recognition, ICPR 2014, 2014, pp. 1764–1769.

[18] L. Rabiner, A tutorial on Hidden Markov Models and selected applications in speech recognition, Proc. of IEEE 77 (2) (1989) 257–286.

[19] G. Grossman, C. Shapiro, Informative advertising with differentiated products, Review of Economic Studies 51 (1984) 63–82.

[20] D. Hallerman, Audience ad targeting: Data and privacy issues (2010).

[21] A. De Corniére, R. Nijs, Online advertising and privacy, RAND Journal of Economics 47 (1) (2016) 4872.

[22] L. Olejnik, C. Castelluccia, A. Janc, Why johnny cant browse in peace: On the uniqueness of web browsing history patterns, in: Proc. of W. on Hot Topics in Privacy Enhancing Technologies, 2012.

[23] P. Lovato, M. Bicego, C. Segalin, A. Perina, N. Sebe, M. Cristani, Faved! biometrics: Tell me which image you like and i'll tell you who you are, IEEE Transactions on Information Forensics and Security 9 (3) (2014) 364–374.

[24] L. Esteban, A. Gil, J. Hernandez, Informative advertising and optimal targeting in a monopoly, Journal of Industrial Economics 49 (2) (2001) 161180.

[25] G. Chicco, O.-M. Ionel, R. Porumb, Electrical load pattern grouping based on centroid model with ant colony clustering, IEEE Transactions on Power Systems 28 (2) (2013) 1706–1715.

[26] S. Orfanidis, Optimum Signal Processing: an Introduction, 2nd Edition, Prentice-Hall, 1996.

[27] M. Vinyals, F. Bistaffa, A. Farinelli, A. Rogers, Coalitional energy purchasing in the smart grid, in: Proc. of IEEE Int. Energy Conference and Exhibition (ENERGYCON 12), 2012, pp. 848 –853.

[28] P. Janovsky, S. DeLoach, Increasing coalition stability in large-scale coalition formation with self-interested agents, in: Proc. European Conf. on Artificial Intelligence, 2016, pp. 1606–1607.

[29] A. Trindade, Uci maching learning repository - electricityloaddiagrams20112014 data set (2014).
URL http://archive.ics.uci.edu/ml/datasets/ElectricityLoadDiagrams20112014

[30] H. Moon, P. Phillips, Computational and performance aspects of pca-based face-recognition algorithms, Perception 30 (2001) 303 – 321.

[31] M. Bicego, V. Murino, M. Figueiredo, A sequential pruning strategy for the selection of the number of states in Hidden Markov Models, Pattern Recognition Letters 24 (9–10) (2003) 1395–1407.

[32] R. Yampolskiy, V. Govindaraju, Behavioural biometrics: a survey and classification, Int. J. Biometrics 1 (1) (2008) 81–113.